GenCore version 5.1.4_p5_4578

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 1, 2003, 08:47:51 ; Search time 64.5 Seconds
(without alignments)
5739.073 Million cell updates/sec

Title: US-09-768-781-2
Perfect score: 2543
Sequence: 1 atgaacacaagaccacaaca.....caaggcaaagtgttgtctga 1389

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09768781/runat_01042003_084729_4923/app_query.fasta_1.1543
-DB=A_Geneseq_101002 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09768781_@CGN_1_1_114_@runat_01042003_084729_4923 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	881	34.6	410	23	ABB97282	Novel human protei
2	660	26.0	131	23	ABB89300	Human polypeptide
3	627	24.7	125	22	AAU86530	Novel human connec
4	614	24.1	216	22	ABB27443	Human peptide #94
5	614	24.1	216	22	ABB32592	Peptide #98 encode
6	614	24.1	216	22	ABB18089	Protein #88 encode
7	614	24.1	216	22	AAM53421	Human brain expres
8	614	24.1	216	22	AAM65799	Human bone marrow
9	614	24.1	216	22	AAM13661	Peptide #95 encode
10	614	24.1	216	22	AAM26060	Peptide #97 encode
11	614	24.1	216	22	AAM01411	Peptide #93 encode
12	614	24.1	216	23	ABG35433	Human peptide enco
13	361	14.2	128	22	ABB22596	Protein #4595 enco
14	361	14.2	128	22	AAM58002	Human brain expres
15	174.5	6.9	86	22	ABB29870	Peptide #2521 enco
16	174.5	6.9	86	22	ABB35054	Peptide #2560 enco
17	174.5	6.9	86	22	ABB20466	Protein #2465 enco
18	174.5	6.9	86	22	AAM55869	Human brain expres
19	174.5	6.9	86	22	AAM68240	Human bone marrow
20	174.5	6.9	86	22	AAM16062	Peptide #2496 enco
21	174.5	6.9	86	22	AAM28558	Peptide #2595 enco
22	174.5	6.9	86	22	AAM03789	Peptide #2471 enco
23	174.5	6.9	86	23	ABG37776	Human peptide enco
24	169	6.6	129	19	AAW81517	XK related Y (XKRY
25	127	5.0	686	22	AAU29285	Human PRO polypept
26	121.5	4.8	783	23	ABB91662	Herbicidally activ
27	119.5	4.7	264	22	AAE03438	Human gene 12 enco
28	119.5	4.7	264	23	ABG63337	Human albumin fusi
29	114	4.5	800	23	ABP40810	Staphylococcus epi
30	113	4.4	264	22	AAE03484	Human gene 12 enco
31	113	4.4	264	23	ABG63335	Human albumin fusi
32	111.5	4.4	785	23	ABB91027	Herbicidally activ
33	108.5	4.3	497	20	AAW95461	L. helveticus pept
34	108.5	4.3	786	22	AAU34510	E. coli cellular p
35	108.5	4.3	858	23	ABP27959	Streptococcus poly
36	108	4.2	751	22	AAU38408	Salmonella typhi c
37	107	4.2	548	23	ABP30418	Streptococcus poly
38	107	4.2	550	23	ABP28802	Streptococcus poly
39	106	4.2	408	20	AAY19788	B. burgdorferi ant
40	106	4.2	440	20	AAY19787	B. burgdorferi ant
41	105.5	4.1	339	23	ABB55349	Lactococcus lactis
42	104	4.1	353	22	ABG07351	Novel human diagno
43	103.5	4.1	290	23	ABP40048	Staphylococcus epi
44	103.5	4.1	663	22	AAB85767	Human seven-transm
45	102	4.0	246	22	ABG10436	Novel human diagno

ALIGNMENTS

RESULT 1
ABB97282
ID ABB97282 standard; Protein; 410 AA.
XX
AC ABB97282;
XX
DT 27-JUN-2002 (first entry)
XX
DE Novel human protein SEQ ID NO: 550.
XX
KW Human; antianaemic; vulnerary; antiinflammatory; immunomodulator;
KW antiinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
KW neuroprotective; antiparkinsonian; protein therapy; EST;
KW expressed sequence tag.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX

PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR N-PSDB; ABN32468.
XX
PT An isolated polynucleotide for treating diseases associated with its
PT encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Example 2; SEQ ID NO 550; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a protein of the invention.
XX
SQ Sequence 410 AA;

Alignment Scores:

Pred. No.:	3.45e-88	Length:	410
Score:	881.00	Matches:	168
Percent Similarity:	65.85%	Conservative:	75
Best Local Similarity:	45.53%	Mismatches:	116
Query Match:	34.64%	Indels:	10
DB:	23	Gaps:	4

US-09-768-781-2 (1-1389) x ABB97282 (1-410)

```
Qy    235 TGGATGACATACACCTTTTCTTTCTTTATGTTTTCATCCATTATGGTCCAGTTGACCCTC 294
          |||         |||      |||   :::           :::|||||||||||||||
Db      2 TrpGlnAlaLeuThrLeuLeuPheSerLeuLeuProCysAlaLeuValGlnLeuThrLeu 21

Qy    295 ATTTTTGTCCACAGAGATCTAGCCAAAGATAAACCGCTATCATTATTTATGCATCTAATC 354
          :::|||||||||||||||||||:::::|||:::||||||   |||   :::||||||:::
Db     22 LeuPheValHisArgAspLeuSerArgAspArgProLeuValLeuLeuLeuHisLeuLeu 41

Qy    355 CTCTTGGGACCTGTTATCAGATGTTTGGAGGCCATGATTAAGTACCTCACACTGTGGAAG 414
             |||||||||:::   ||||||   |||            |||         ::::::
Db     42 GlnLeuGlyProLeuPheArgCysPheGluValPheCysIleTyr---------PheGln 58

Qy    415 AAAGAGGAGCAGGAGGAGCCCTATGTCAGCCTCACCCGAAAGAAG---ATGCTAATAGAT 471
                   |||||||||||||||||:::|||:::|||:::   |||      :::
Db     59 SerGlyAsnAsnGluGluProTyrValSerIleThrLysLysArgGlnMetProLysAsn 78

Qy    472 GGCGAGGAGGTGCTGATAGAATGGGAGGTGGGCCACTCCATCCGGACCCTGGCTATGCAC 531
          |||           ||||||   |||||||||   :::          |||      |||
Db     79 GlyLeuSerGluGluIleGluLysGluValGlyGlnAlaGluGlyLysLeuIleThrHis 98

Qy    532 CGCAATGCCTACAAACGTATGTCACAGATCCAAGCCTTCCTGGGCTCAGTGCCCCAGCTG 591
          |||:::|||:::   |||   |||   ||||||||||||||||||||   |||||||||
Db     99 ArgSerAlaPheSerArgAlaSerValIleGlnAlaPheLeuGlySerAlaProGlnLeu 118

Qy    592 ACCTATCAGCTCTATGTGAGCCTGATCTCTGCAGAGGTTCCCCTGGGTAGAGTTGTGCTA 651
          |||   |||||||||:::|||::::::     :::|||   :::||||||   :::|||
Db    119 ThrLeuGlnLeuTyrIleSerValMetGlnGlnAspValThrValGlyArgSerLeuLeu 138

Qy    652 ATGGTATTTTCCCTGGTATCTGTCACCTATGGGGCCACCCTTTGCAATATGTTGGCTATC 711
          |||      ||||||:::|||:::   |||||||||      ||||||:::|||||||||
Db    139 MetThrIleSerLeuLeuSerIleValTyrGlyAlaLeuArgCysAsnIleLeuAlaIle 158

Qy    712 CAGATCAAGTACGATGACTACAAGATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACC 771
          :::|||||||||||||:::|||::::::::::::   ||||||      :::||||||
Db    159 LysIleLysTyrAspGluTyrGluValLysValLysProLeuAlaTyrValCysIlePhe 178

Qy    772 ATCTGGCGGACATTGGAGATCACTTCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACT 831
          :::||||||:::   ||||||   :::|||:::::::::|||||||||||||:::::::
Db    179 LeuTrpArgSerPheGluIleAlaThrArgValValValLeuValLeuPheThrSerVal 198

Qy    832 TTGAAATTGAAGGCTGTGCCCTTCCTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCC 891
          ||||||          |||      :::::::::||||||                  |||
Db    199 LeuLysThrTrpValValValIleIleLeuIleAsnPhePheSerPhePheLeuTyrPro 218

Qy    892 TGGATTAAGTTCTGGAGAAGTGGTGCCCAGATGCCCAATAACATTGAGAAAAACTTCAGC 951
          ||||||   ||||||   ||||||:::      |||   ||||||||||||      |||
Db    219 TrpIleLeuPheTrpCysSerGlySerProPheProGluAsnIleGluLysAlaLeuSer 238

Qy    952 CGGGTCGGCACTCTGGTGGTCCTGATTTCAGTCACCATCCTCTATGCTGGCATCAACTTC 1011
          ||||||||||||   :::||||||      :::|||:::||||||   |||||||||
Db    239 ArgValGlyThrThrIleValLeuCysPheLeuThrLeuLeuTyrThrGlyIleAsnMet 258

Qy   1012 TCTTGCTGGTCAGCTTTGCAGTTGAGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAG 1071
             ||||||||||||:::|||||:::::          ||||||:::   |||
Db    259 PheCysTrpSerAlaValGlnLeuLysIleAspSerProAspLeuIleSerLysSerHis 278

Qy   1072 AACTGGGGACATATGGGCCTGCACTATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTC 1131
          ||||||      :::   ::::::|||   :::|||   :::||||||   |||:::::::
Db    279 AsnTrpTyrGlnLeuLeuValTyrTyrMetIleArgPheIleGluAsnAlaIleLeuLeu 298

Qy   1132 TTGGTTTTTAAGTTCTTTGGAGTGAAAGTGTTACTGAATTACTGTCATTCCTTGATTGCC 1191
          |||::::::      |||          :::   :::      |||      |||:::
Db    299 LeuLeuTrpTyrLeuPheLysThrAspIleTyrMetTyrValCysAlaProLeuLeuVal 318

Qy   1192 TTGCAGCTCATTATTGCTTATCTGATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTAC 1251
          |||||||||:::|||   |||      :::|||   |||||||||:::|||:::|||:::
Db    319 LeuGlnLeuLeuIleGlyTyrCysThrAlaIleLeuPheMetLeuValPheTyrGlnPhe 338

Qy   1252 TTGCATCCATTGCGCTCACTCTTCACCCATAATGTAGTAGAC------------TACCTC 1299
             ||||||   :::   ||||||:::   :::|||   :::            :::|||
Db    339 PheHisProCysLysLysLeuPheSerSerSerValSerGluGlyPheGlnArgTrpLeu 358

Qy   1300 CATTGTGTCTGC------TGTCACCAG 1320
             |||   |||      |||   |||
Db    359 ArgCysPheCysTrpAlaCysArgGln 367
```

RESULT 2
ABB89300
ID ABB89300 standard; Protein; 131 AA.
XX
AC ABB89300;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 1676.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001WO-US16450.
XX
PR 19-MAY-2000; 2000US-205515P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX

PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
DR N-PSDB; ABL89709.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
PS Claim 11; SEQ ID NO 1676; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 131 AA;

Alignment Scores:

Pred. No.:	6.01e-64	Length:	131
Score:	660.00	Matches:	125
Percent Similarity:	96.18%	Conservative:	1
Best Local Similarity:	95.42%	Mismatches:	5
Query Match:	25.95%	Indels:	0
DB:	23	Gaps:	0

US-09-768-781-2 (1-1389) x ABB89300 (1-131)

```
Qy    922 ATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTGATTTCA 981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetProAsnAsnIleGluLysAsnPheSerArgValGlyThrLeuValValLeuIleSer 20

Qy    982 GTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTGAGGTTG 1041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 ValThrIleLeuTyrAlaGlyIleAsnPheSerCysTrpSerAlaLeuGlnLeuArg*** 40

Qy   1042 GCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCACTATAGT 1101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 AlaAspArgAspLeuValAspLysGlyGlnAsnTrpGlyHisMetGlyLeuHisTyrSer 60

Qy   1102 GTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTGAAAGTG 1161
          |||:::|||||||||||||||||||||||||||||||||||   ||||||||||||
Db     61 ValLysLeuValGluAsnValIleMetValLeuValPheLysPhe***GlyValLysVal 80

Qy   1162 TTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTGATTTCC 1221
            |||||||||||||||   |||   ||||||||||||||||||||||||||||||||
Db     81 ***LeuAsnTyrCysHis***Leu***AlaLeuGlnLeuIleIleAlaTyrLeuIleSer 100

Qy   1222 ATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCACTCTTCACCCAT 1281
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 IleGlyPheMetLeuLeuPhePheGlnTyrLeuHisProLeuArgSerLeuPheThrHis 120

Qy   1282 AATGTAGTAGACTACCTCCATTGTGTCTGCTGT 1314
          |||||||||||||||||||||||||||||||
Db    121 AsnValValAspTyrLeuHisCysValCysCys 131
```

RESULT 3
AAU86530

ID AAU86530 standard; Protein; 125 AA.
XX
AC AAU86530;
XX
DT 21-MAY-2002 (first entry)
XX
DE Novel human connective tissue related polypeptide #96.
XX
KW Human; connective tissue related disorder; cancer; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200155343-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01322.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-565190/63.
DR N-PSDB; ABK41708.
XX
PT Nucleic acid encoding novel connective tissue associated polypeptides,
PT used in diagnosing, preventing, treating or ameliorating a disorder
PT such as cancer or rheumatoid arthritis -
XX
PS Claim 11; SEQ ID No 595; 673pp; English.
XX
CC The present invention relates to the isolation of novel human
CC connective tissue related polypeptides and the polynucleotide
CC (cDNA and genomic) sequences encoding them. The sequences of the
CC invention are useful in the diagnosis, treatment, prevention and/or
CC prognosis of diseases associated with connective tissue(s), including
CC cancer. The polynucleotide sequences of the invention are also useful
CC in gene therapy. AAU86435-AAU86923 represent the novel human connective
CC tissue related polypeptides.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 125 AA;

Alignment Scores:

Pred. No.:	2.64e-60	Length:	125
Score:	627.00	Matches:	120
Percent Similarity:	96.03%	Conservative:	1
Best Local Similarity:	95.24%	Mismatches:	4
Query Match:	24.66%	Indels:	1
DB:	22	Gaps:	0

US-09-768-781-2 (1-1389) x AAU86530 (1-125)

```
Qy    939 GAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTGATTTCAGTCACCATCCTCTATGC 998
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db      1 GluLysLeuGlnProGlyArgHisSerGlyGly-LeuIleSerValThrIleLeuTyrAl 20

Qy    999 TGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTGAGGTTGGCAGACAGAGATCTCGT 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     20 aGlyIleAsnPheSerCysTrpSerAlaLeuGlnLeuArgLeuAlaAspArgAspLeuVa 40

Qy   1059 CGACAAAGGGCAGAACTGGGGACATATGGGCCTGCACTATAGTGTGAGGTTGGTAGAGAA 1118
          ||||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db     40 lAspLysGlyGlnAsnTrpGlyHisMetGlyLeuHisTyrSerValLysLeuValGluAs 60

Qy   1119 TGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTGAAAGTGTTACTGAATTACTGTCA 1178
          |||||||||||||||||||||||||||    ||||||||||||   |||||||||||||||
Db     60 nValIleMetValLeuValPheLysPhe***GlyValLysVal***LeuAsnTyrCysHi 80

Qy   1179 TTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTGATTTCCATTGGCTTCATGCTCCT 1238
          |   |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 s***Leu***AlaLeuGlnLeuIleIleAlaTyrLeuIleSerIleGlyPheMetLeuLe 100

Qy   1239 TTTCTTCCAGTACTTGCATCCATTGCGCTCACTCTTCACCCATAATGTAGTAGACTACCT 1298
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    100 uPhePheGlnTyrLeuHisProLeuArgSerLeuPheThrHisAsnValValAspTyrLe 120

Qy   1299 CCATTGTGTCTGCTGT 1314
          ||||||||||||||||
Db    120 uHisCysValCysCys 125
```

```
RESULT 4
ABB27443
ID   ABB27443 standard; Peptide; 216 AA.
XX
AC   ABB27443;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human peptide #94 encoded by breast cell single exon nucleic acid probe.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 27; SEQ ID NO 10411; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a peptide encoded by a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59        Length:         216
Score:                  614.00          Matches:        110
Percent Similarity:     76.77%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     46
Query Match:            24.14%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x ABB27443 (1-216)

Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::|||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||       |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||:::::   |||       |||   |||:::||||||||||||:::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||       ||||||::::::|||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   |||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
                :::|||:::|||||||||   ||||||||||||||||||||||||:::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   ||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::|||||||||||||||:::||||||||::||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   ||||||||||||:::||||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   ||||||||||||||||:::||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198

RESULT 5
ABB32592
ID   ABB32592 standard; Peptide; 216 AA.
XX
AC   ABB32592;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Peptide #98 encoded by human foetal liver single exon nucleic acid probe.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
```

```
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 27; SEQ ID NO 25227; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a peptide encoded by a single exon
CC   nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59        Length:         216
Score:                  614.00          Matches:        110
Percent Similarity:     76.77%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     46
Query Match:            24.14%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x ABB32592 (1-216)

Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||::::::   |||      |||   |||:::||||||||||||::::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||      ||||||::::::||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   |||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
              :::|||:::|||||||||   |||||||||||||||||||||||::::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   ||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::||||||||||||||::|||||||||:::|||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   ||||||||||||:::||||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   ||||||||||:|||||||:::|||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198

RESULT 6
ABB18089
ID   ABB18089 standard; Protein; 216 AA.
XX
AC   ABB18089;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Protein #88 encoded by probe for measuring heart cell gene expression.
XX
KW   Human; gene expression; heart; microarray; vascular system;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 15; SEQ ID No 19859; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart (see
CC   ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC   probe. The probes may be used for predicting, measuring and displaying
CC   gene expression in samples derived from the human heart via microarrays.
CC   By measuring gene expression, the probes are useful for predicting,
CC   diagnosing, grading, staging, monitoring and prognosing diseases of the
CC   human heart and vascular system e.g. cardiovascular disease,
CC   hypertension, cardiac arrhythmias and congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59        Length:         216
Score:                  614.00          Matches:        110
Percent Similarity:     76.77%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     46
Query Match:            24.14%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x ABB18089 (1-216)

Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||::::::   |||      |||   |||:::||||||||||||::::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60
```

```
Qy    856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
          |||::::::   :::   :::   |||       ||||||::::::||||||:::||||||
Db     61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy    916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
          |||   :::|||   |||   |||   |||   :::   |||||||||:::::::::|||
Db     81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy    976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
             :::|||:::|||||||||   |||||||||||||||||||||||:::::|||
Db    101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy   1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
          :::|||:::|||       ::::::|||       |||   |||||||||       ||||||
Db    121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy   1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
          ||||||   :::   :::||||||||||||||||:::||||||||||:::|||||||||
Db    141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy   1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
          |||   |||||||||   |||   |||||||||||||:::|||||||||||||:::||||||
Db    161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy   1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
          ::::::   ||||||||||||||||||:::||||||||||:::|||   :::|||
Db    181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198
```

RESULT 7
AAM53421
ID AAM53421 standard; Protein; 216 AA.
XX
AC AAM53421;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 25526.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 25526; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 216 AA;

Alignment Scores:

Pred. No.:	9.24e-59	Length:	216
Score:	614.00	Matches:	110
Percent Similarity:	76.77%	Conservative:	42
Best Local Similarity:	55.56%	Mismatches:	46
Query Match:	24.14%	Indels:	0
DB:	22	Gaps:	0

US-09-768-781-2 (1-1389) x AAM53421 (1-216)

```
Qy    676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
          ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db      1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy    736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
          |||:::|||   |||:::|||       |||:::   :::||||||   ||||||:::
Db     21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy    796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
          ||||||::::::   |||       |||   |||:::||||||||||||::::::|||
Db     41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy    856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
          |||::::::   :::   :::   |||       ||||||:::::||||||:::||||||
Db     61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy    916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
          |||   :::|||   |||   |||   |||   :::   |||||||||:::::::::|||
Db     81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy    976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
             :::|||:::|||||||||   |||||||||||||||||||||||:::::|||
Db    101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy   1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
          :::|||:::|||       ::::::|||       |||   |||||||||       ||||||
Db    121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy   1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
          ||||||   :::   :::||||||||||||||||:::||||||||||:::|||||||||
Db    141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy   1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
          |||   |||||||||   |||   |||||||||||||:::|||||||||||||:::||||||
Db    161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy   1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
          ::::::   ||||||||||||||||||:::||||||||||:::|||   :::|||
Db    181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198
```

RESULT 8
AAM65799
ID AAM65799 standard; Protein; 216 AA.
XX
AC AAM65799;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 26105.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.

```
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 26105; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is a
CC   protein encoded by one of the probes of the invention.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59        Length:         216
Score:                  614.00          Matches:        110
Percent Similarity:     76.77%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     46
Query Match:            24.14%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x AAM65799 (1-216)

Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||::::::   |||      |||   |||:::||||||||||||::::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||      ||||||::::::||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   ||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
              :::|||:::|||||||||   |||||||||||||||||||||||:::::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   |||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::||||||||||||||:::||||||||:::||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   |||||||||||:::|||||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   |||||||||||||||||:::|||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198

RESULT 9
AAM13661
ID   AAM13661 standard; Protein; 216 AA.
XX
AC   AAM13661;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Peptide #95 encoded by probe for measuring cervical gene expression.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 27; SEQ ID No 18487; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
CC   by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC   can be used to produce a single exon microarray, which can be used for
CC   measuring human gene expression in a sample derived from human cervical
CC   epithelial cells. By measuring gene expression, the probes are therefore
CC   useful in grading and/or staging of diseases of the cervix, notably
CC   cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59        Length:         216
Score:                  614.00          Matches:        110
Percent Similarity:     76.77%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     46
Query Match:            24.14%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x AAM13661 (1-216)
```

```
Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::|||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||::::::   |||      |||   |||:::||||||||||||::::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||      ||||||::::::||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   ||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
              :::|||:::|||||||||   ||||||||||||||||||||||||:::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   |||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::||||||||||||||:::|||||||||:::|||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   ||||||||||:::|||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   ||||||||||||||||||:::|||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198
```

RESULT 10
AAM26060
ID AAM26060 standard; Protein; 216 AA.
XX
AC AAM26060;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #97 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
PS Claim 27; SEQ ID No 26329; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 216 AA;

Alignment Scores:

Pred. No.:	9.24e-59	Length:	216
Score:	614.00	Matches:	110
Percent Similarity:	76.77%	Conservative:	42
Best Local Similarity:	55.56%	Mismatches:	46
Query Match:	24.14%	Indels:	0
DB:	22	Gaps:	0

US-09-768-781-2 (1-1389) x AAM26060 (1-216)

```
Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::|||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||::::::   |||      |||   |||:::||||||||||||::::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||      ||||||::::::||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   ||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
              :::|||:::|||||||||   ||||||||||||||||||||||||:::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   |||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::||||||||||||||:::|||||||||:::|||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   ||||||||||:::|||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   ||||||||||||||||||:::|||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198
```

RESULT 11
AAM01411
ID AAM01411 standard; Protein; 216 AA.
XX
AC AAM01411;
XX

DT 09-OCT-2001 (first entry)
XX
DE Peptide #93 encoded by probe for measuring human breast gene expression.
XX
KW Probe; human; breast disease; breast cancer; development disorder;
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US00661.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast -
XX
PS Claim 27; SEQ ID No 10151; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes
CC (see AAI00010-AAI10067). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for measuring human gene expression in
CC a human breast sample, where the probe hybridises at high stringency to a
CC nucleic acid expressed in the human breast. The probes are useful for
CC predicting, diagnosing, grading, staging, monitoring and prognosing
CC diseases of the human breast, particularly those diseases with polygenic
CC aetiology. The diseases include: breast cancer, disorders of development,
CC inflammatory diseases of the breast, fibrocystic changes, proliferative
CC breast disease and non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 216 AA;

Alignment Scores:

Pred. No.:	9.24e-59	Length:	216
Score:	614.00	Matches:	110
Percent Similarity:	76.77%	Conservative:	42
Best Local Similarity:	55.56%	Mismatches:	46
Query Match:	24.14%	Indels:	0
DB:	22	Gaps:	0

US-09-768-781-2 (1-1389) x AAM01411 (1-216)

```
Qy    676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
          ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db      1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy    736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
          |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db     21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy    796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
          ||||||::::::   |||      |||   |||:::||||||||||||:::::::|||
Db     41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy    856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
          |||::::::   :::   :::   |||      ||||||::::::||||||:::||||||
Db     61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy    916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
          |||   :::|||   |||   |||   |||   :::   |||||||||:::::::|||
Db     81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy    976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
          :::|||:::|||||||||   ||||||||||||||||||||||:::::|||
Db    101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy   1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
          :::|||:::|||      ::::::|||      |||   |||||||      ||||||
Db    121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy   1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
          ||||||   :::   :::|||||||||||||||:::||||||||:::|||||||||
Db    141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPhePheGlyGly 160

Qy   1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
          |||   |||||||||   |||   |||||||||||:::|||||||||||:::||||||
Db    161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy   1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
          ::::::   ||||||||||||||||||:::|||||||||:::|||   :::|||
Db    181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198
```

RESULT 12
ABG35433
ID ABG35433 standard; Peptide; 216 AA.
XX
AC ABG35433;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 25098.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesis; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -
XX
PS Claim 27; SEQ ID No 25098; 634pp; English.

```
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a peptide/protein
CC   encoded by a single exon probe of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   216 AA;

Alignment Scores:
Pred. No.:              9.24e-59      Length:         216
Score:                  614.00        Matches:        110
Percent Similarity:     76.77%        Conservative:   42
Best Local Similarity:  55.56%        Mismatches:     46
Query Match:            24.14%        Indels:         0
DB:                     23            Gaps:           0

US-09-768-781-2 (1-1389) x ABG35433 (1-216)

Qy     676 ACCTATGGGGCCACCCTTTGCAATATGTTGGCTATCCAGATCAAGTACGATGACTACAAG 735
           ||||||||||||      ||||||:::||||||||||||||||      ||||||
Db       1 ThrTyrGlyAlaIleArgCysAsnIleLeuAlaIleGlnIleSerAsnAspAspThrThr 20

Qy     736 ATTCGCCTTGGGCCACTAGAAGTCCTCTGCATCACCATCTGGCGGACATTGGAGATCACT 795
           |||:::|||   |||:::|||      |||:::   :::||||||   ||||||:::
Db      21 IleLysLeuProProIleGluPhePheCysValValMetTrpArgPheLeuGluValIle 40

Qy     796 TCCCGCCTCCTGATTCTGGTGCTCTTCTCAGCCACTTTGAAATTGAAGGCTGTGCCCTTC 855
           ||||||:::::   |||      |||   |||:::||||||||||||:::::|||
Db      41 SerArgValValThrLeuAlaPhePheIleAlaSerLeuLysLeuLysSerLeuProVal 60

Qy     856 CTAGTGCTCAACTTCCTGATCATCCTCTTTGAGCCCTGGATTAAGTTCTGGAGAAGTGGT 915
           |||::::::   :::   :::   |||      ||||||:::::||||||:::||||||
Db      61 LeuLeuIleIleTyrPheValSerLeuLeuAlaProTrpLeuGluPheTrpLysSerGly 80

Qy     916 GCCCAGATGCCCAATAACATTGAGAAAAACTTCAGCCGGGTCGGCACTCTGGTGGTCCTG 975
           |||   :::|||   |||   |||   |||   :::   |||||||||:::::::::|||
Db      81 AlaHisLeuProGlyAsnLysGluAsnAsnSerAsnMetValGlyThrValLeuMetLeu 100

Qy     976 ATTTCAGTCACCATCCTCTATGCTGGCATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTG 1035
           :::|||:::|||||||||   ||||||||||||||||||||||||:::::|||
Db     101 PheLeuIleThrLeuLeuTyrAlaAlaIleAsnPheSerCysTrpSerAlaValLysLeu 120

Qy    1036 AGGTTGGCAGACAGAGATCTCGTCGACAAAGGGCAGAACTGGGGACATATGGGCCTGCAC 1095
           :::|||:::|||      ::::::|||      |||   |||||||||      ||||||
Db     121 GlnLeuSerAspAspLysIleIleAspGlyArgGlnArgTrpGlyHisArgIleLeuHis 140

Qy    1096 TATAGTGTGAGGTTGGTAGAGAATGTGATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTG 1155
           ||||||   :::   :::|||||||||||||||:::|||||||||:::|||||||||
Db     141 TyrSerPheGlnPheLeuGluAsnValIleMetIleLeuValPheArgPheGlyGly 160

Qy    1156 AAAGTGTTACTGAATTACTGTCATTCCTTGATTGCCTTGCAGCTCATTATTGCTTATCTG 1215
           |||   |||||||||   |||   |||||||||||:::|||||||||||:::||||||
Db     161 LysThrLeuLeuAsnCysCysAspSerLeuIleAlaValGlnLeuIleIleSerTyrLeu 180

Qy    1216 ATTTCCATTGGCTTCATGCTCCTTTTCTTCCAGTACTTGCATCCATTGCGCTCA 1269
           ::::::   |||||||||||||||||:::|||||||||:::|||   :::|||
Db     181 LeuAlaThrGlyPheMetLeuLeuPheTyrGlnTyrLeuTyrProTrpGlnSer 198

RESULT 13
ABB22596
ID   ABB22596 standard; Protein; 128 AA.
XX
AC   ABB22596;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Protein #4595 encoded by probe for measuring heart cell gene expression.
XX
KW   Human; gene expression; heart; microarray; vascular system;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 15; SEQ ID No 24366; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart (see
CC   ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC   probe. The probes may be used for predicting, measuring and displaying
CC   gene expression in samples derived from the human heart via microarrays.
CC   By measuring gene expression, the probes are useful for predicting,
CC   diagnosing, grading, staging, monitoring and prognosing diseases of the
CC   human heart and vascular system e.g. cardiovascular disease,
CC   hypertension, cardiac arrhythmias and congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
```

```
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   128 AA;

Alignment Scores:
Pred. No.:              7.34e-31        Length:         128
Score:                  361.00          Matches:        66
Percent Similarity:     81.65%          Conservative:   23
Best Local Similarity:  60.55%          Mismatches:     20
Query Match:            14.20%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x ABB22596 (1-128)

Qy    943 AACTTCAGCCGGGTCGGCACTCTGGTGGTCCTGATTTCAGTCACCATCCTCTATGCTGGC 1002
          |||   :::   |||||||||:::::::::|||      :::|||:::|||||||||
Db      2 AsnSerAsnMetValGlyThrValLeuMetLeuPheLeuIleThrLeuLeuTyrAlaAla 21

Qy   1003 ATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTGAGGTTGGCAGACAGAGATCTCGTCGAC 1062
          ||||||||||||||||||||||:::::::|||:::|||:::|||      ::::::|||
Db     22 IleAsnPheSerCysTrpSerAlaValLysLeuGlnLeuSerAspAspLysIleIleAsp 41

Qy   1063 AAAGGGCAGAACTGGGGACATATGGGCCTGCACTATAGTGTGAGGTTGGTAGAGAATGTG 1122
               |||   |||||||||      ||||||||||||   :::   :::|||||||||
Db     42 GlyArgGlnArgTrpGlyHisArgIleLeuHisTyrSerPheGlnPheLeuGluAsnVal 61

Qy   1123 ATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTGAAAGTGTTACTGAATTACTGTCATTCC 1182
          ||||||:::|||||||||:::|||||||||   |||   |||||||||   |||   |||
Db     62 IleMetIleLeuValPheArgPhePheGlyGlyLysThrLeuLeuAsnCysCysAspSer 81

Qy   1183 TTGATTGCCTTGCAGCTCATTATTGCTTATCTGATTTCCATTGGCTTCATGCTCCTTTTC 1242
          |||||||||:::||||||||||:::||||||::::::   ||||||||||||||||||
Db     82 LeuIleAlaValGlnLeuIleIleSerTyrLeuLeuAlaThrGlyPheMetLeuLeuPhe 101

Qy   1243 TTCCAGTACTTGCATCCATTGCGCTCA 1269
          :::|||||||||:::|||   :::|||
Db    102 TyrGlnTyrLeuTyrProTrpGlnSer 110

RESULT 14
AAM58002
ID   AAM58002 standard; Protein; 128 AA.
XX
AC   AAM58002;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe encoded protein SEQ ID NO: 30107.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 30107; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is a protein encoded by one of
CC   the probes of the invention.
XX
SQ   Sequence   128 AA;

Alignment Scores:
Pred. No.:              7.34e-31        Length:         128
Score:                  361.00          Matches:        66
Percent Similarity:     81.65%          Conservative:   23
Best Local Similarity:  60.55%          Mismatches:     20
Query Match:            14.20%          Indels:         0
DB:                     22              Gaps:           0

US-09-768-781-2 (1-1389) x AAM58002 (1-128)

Qy    943 AACTTCAGCCGGGTCGGCACTCTGGTGGTCCTGATTTCAGTCACCATCCTCTATGCTGGC 1002
          |||   :::   |||||||||:::::::::|||      :::|||:::|||||||||
Db      2 AsnSerAsnMetValGlyThrValLeuMetLeuPheLeuIleThrLeuLeuTyrAlaAla 21

Qy   1003 ATCAACTTCTCTTGCTGGTCAGCTTTGCAGTTGAGGTTGGCAGACAGAGATCTCGTCGAC 1062
          ||||||||||||||||||||||:::::::|||:::|||:::|||      ::::::|||
Db     22 IleAsnPheSerCysTrpSerAlaValLysLeuGlnLeuSerAspAspLysIleIleAsp 41

Qy   1063 AAAGGGCAGAACTGGGGACATATGGGCCTGCACTATAGTGTGAGGTTGGTAGAGAATGTG 1122
               |||   |||||||||      ||||||||||||   :::   :::|||||||||
Db     42 GlyArgGlnArgTrpGlyHisArgIleLeuHisTyrSerPheGlnPheLeuGluAsnVal 61

Qy   1123 ATCATGGTCTTGGTTTTTAAGTTCTTTGGAGTGAAAGTGTTACTGAATTACTGTCATTCC 1182
          ||||||:::|||||||||:::|||||||||   |||   |||||||||   |||   |||
Db     62 IleMetIleLeuValPheArgPhePheGlyGlyLysThrLeuLeuAsnCysCysAspSer 81

Qy   1183 TTGATTGCCTTGCAGCTCATTATTGCTTATCTGATTTCCATTGGCTTCATGCTCCTTTTC 1242
          |||||||||:::||||||||||:::||||||::::::   ||||||||||||||||||
Db     82 LeuIleAlaValGlnLeuIleIleSerTyrLeuLeuAlaThrGlyPheMetLeuLeuPhe 101

Qy   1243 TTCCAGTACTTGCATCCATTGCGCTCA 1269
          :::|||||||||:::|||   :::|||
Db    102 TyrGlnTyrLeuTyrProTrpGlnSer 110

RESULT 15
ABB29870
ID   ABB29870 standard; Peptide; 86 AA.
XX
AC   ABB29870;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Peptide #2521 encoded by breast cell single exon nucleic acid probe.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
```

PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 27; SEQ ID NO 12838; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 86 AA;

Alignment Scores:

Pred. No.:	2.69e-10	Length:	86
Score:	174.50	Matches:	39
Percent Similarity:	69.86%	Conservative:	12
Best Local Similarity:	53.42%	Mismatches:	21
Query Match:	6.86%	Indels:	1
DB:	22	Gaps:	1

US-09-768-781-2 (1-1389) x ABB29870 (1-86)

```
Qy     427 GAGGAGCCCTATGTCAGCCTCACCCGAAAGAAG---ATGCTAATAGATGGCGAGGAGGTG 483
           ||||||||||||||||||:::|||:::|||:::   |||      :::|||
Db      14 GluGluProTyrValSerIleThrLysLysArgGlnMetProLysAsnGlyLeuSerGlu 33

Qy     484 CTGATAGAATGGGAGGTGGGCCACTCCATCCGGACCCTGGCTATGCACCGCAATGCCTAC 543
           ||||||  |||||||||  :::          |||      ||||||:::|||:::
Db      34 GluIleGluLysGluValGlyGlnAlaGluGlyLysLeuIleThrHisArgSerAlaPhe 53

Qy     544 AAACGTATGTCACAGATCCAAGCCTTCCTGGGCTCAGTGCCCCAGCTGACCTATCAGCTC 603
           |||  |||  ||||||||||||||||||||  ||||||||||  ||||||
Db      54 SerArgAlaSerValIleGlnAlaPheLeuGlySerAlaProGlnLeuThrLeuGlnLeu 73

Qy     604 TATGTGAGCCTGATCTCTGCAGAGGTTCCCCTGGGTAGA 642
           |||:::|||::::::      :::|||   :::||||||
Db      74 TyrIleSerValMetGlnGlnAspValThrValGlyArg 86
```

Search completed: April 1, 2003, 08:53:20
Job time : 69.5 secs